A TRADITIONAL FRAMEWORK FOR BIKE SHARING SYSTEM BASED ON CLIENT BEHAVIOR ANALYSIS

U. Mohan Srinivas, *Assoc. Prof., Dept. of CSE & MCA, QIS College of Engineering and Technology, Ongole,*

Ch. Rama Krishna, *Final Year Student of M C A, QIS College of Engineering and Technology, Ongole*

Abstract— *The fast improvement of bike sharing frameworks has brought individuals huge accommodation amid the previous decade. Then again, high transport adaptability offers ascend to issues for the two clients and administrators. For clients, dynamic appropriation of shared bikes brought about by uneven client request frequently prompts the look at in or check administration inaccessible at certain stations. For operators, unbalanced bike use accompanies more bike broken and developing upkeep cost. In this paper, we consider upgrading client encounters and rebalance bike usage by guiding clients to various stations with a higher achievement rate of rental and return. Out of the blue, we devise a trip advisor that suggests bike registration and registration stations with joint thought of administration quality and bike use. To guarantee administration quality, we initially foresee the client request of each station to acquire the achievement rate of rental and return later on. Tests demonstrate that the accuracy of our technique is as much as 0.826, which has raised by 25.9% as contrasted and that of the chronicled normal strategy. To rebalance bike utilization, from recorded information, we recognize that one-sided bike use is established from encircled bike flow among couple of dynamic stations. Hence, with characterized station liveliness, we enhance the bike flow by driving clients to move bikes between exceedingly dynamic stations and idle ones.*

Index Terms— Bike Sharing, Bike Registration, Trip advisor

I. INTRODUCTION

With the development of the economy, pollution and destruction caused by human activities to the natural environment were becoming more and more severe in recent years [1], and sustainable development has therefore become a consensus of the international community. In this circumstance [2], bike-sharing systems (BSS) are developed as a replacement for short vehicle journeys due to its low pollution, low energy consumption and high flexibility [3][4]. In addition to the reduction of need for personal vehicle trips, public bike-sharing systems can not only extend the reach of transit and walking trips, providing people with a healthy transportation option [5], but also trigger greater interest in cycling, and increase cycling ridership. By the end of 2016, over 1,100 cities actively [6] operate automated bike-sharing systems deploying an estimate of 2,000,000 public bicycles worldwide. With bike-sharing systems, a user can easily rent a bike with a smart card at a nearby station and return it at another station. However, the advantages cannot cover up the increasingly prominent issues [7][8]. For stations, the user demand is ever-changing and unbalanced, which often leads to the check in or check out service unavailable at some stations and has a negative impact on user experience. For bikes, the usage frequency of each bike is unevenly distributed, posing a problem for both riders and system operators [9][10]. On the one hand, due to the high flexibility of bike sharing system, the system typically ends up with an uneven distribution of bikes across [11][12], the different stations (due to the uncontrolled, uneven demand), often rendering the check in or check out service unavailable at some stations where bicycle docks are either fully occupied or empty [13]. During peak periods, user demand characteristics differ among stations in certain areas. For example, rental demand usually gets larger in workday morning near residential areas, whereas return demand gets larger near commercial districts. At present, operators perform bike redistribution based on monitor video and user complaints [14]. However, this method has exposed the serious lag [15]. It is usually when service unavailable events occur that operators start to give some scheduling instructions. When the vehicle arrives, service unavailable [16], events may have passed for some time, which makes it difficult to meet the needs of users at rush hour. To increase service availability and enhance user experience, studies have been conducted to improve these bike redistribution strategies based on bicycle mobility models and predictions. Most of the previous work focuses on bike usage patterns and rental volume forecasts for each station without considering online information [17]. Less attention has been devoted to demand prediction of each cluster from the view of bike flow mobility patterns which may not fit for recommending stations for users. In conclusion, developing a fine-grained prediction model involving multiple factors has proven to be elusive, and has remained a largely unstudied problem. The main technical [18], challenge is that bike traffic is not only highly dynamic and intercorrelated in both the temporal and spatial domains, but also further influenced by complex issues such as timing and meteorology. To alleviate the unbalanced demand problem, we establish a fine-grained demand forecasting model and predict check in and check out demand on a per-station [19], basis with sub-hour granularity by using random forest algorithm. In our model, offline features such as time and weather are selected to capture the periodic [20], patterns of user demand. Online feature is to reflect the real-time availability of the station which is helpful for abnormal traffic. On the other hand, a small part of bikes is used much more frequently than others [21]. Bikes that are used too much are vulnerable and hence increase repair bills and lead to potential denied service. The very first [22], bicycle from Hangzhou BSS is reported to be rented for over 6,000 times and ridden for more than 20,000 kilometers in 3 years. Similarly, the most tireless bicycle from 2016 has been rented for 5,616 times, over 15 times on average each day. According to Hangzhou [23], public bike-sharing company, the average life of their bicycles is less than 4 years due to longtime high load operation and lack of timely renewal and maintenance. On the contrary, the average life of private bicycles is 10 years and above. Meanwhile, the cost of repair [24], and labor accounts for a large proportion in the overall operating expenses. In 2012, the repair cost of Hangzhou bike-sharing system was near 6 million yuan. In Washington, D.C., the annual maintenance cost was $200 to $300 per bike in the year of 2012.

II. DATA PREPROCESSING

A) Dataset Description:

The Chinese city of Hangzhou has the world's largest public BSS with more than 3300 stations and over 84,000 shared bicycles [11]. Since deployed in May 2008, thousands of bicycles have been rented for more than 700 million times. The concept of public bicycles has since spread to 30 other provinces in China and around 175 cities nationwide [26].

TABLE 1: Primary Fields of Bike sharing system

user id	rent netid	tran date	tran time
8601940	9926	20150601	070641
return netid	return date	return time	bike id
9205	20150601	071635	1708133

The system is classified as a third-generation bike-sharing program due to its IT-based system, automated check-in and check-out, and distinguishable bicycles and docking stations [12]. The dataset used in this paper was collected in June 2015 from our partner who is running Hangzhou BSS. It contains 58,647 bikes and 3,329 stations. Each bike-sharing trip contains an origin and a destination with information of locations and timestamps. The primary fields of the dataset are shown in Table 1. The meteorology dataset contains weather conditions of Hangzhou with totally 48 365 = 17; 520 records. Meteorological observations were updated every half hour and the data format of each record is shown in Table 2.

B) Data cleaning:

The data in the real world are generally incomplete and inconsistent dirty data, so data analysis cannot be directly conducted. Before analyzing the data, it is necessary to perform appropriate data cleaning to obtain high quality data and necessary information.

TABLE 2: Fields in the Meteorology Data set

Time (CST)	Temp (°F)	Dew Point (°F)
12:30 PM	100.4	69.8
Pressure (in)	Humidity (%)	Visibility (mi)
29.65	37	6.2
Wind Dir	Wind Speed (mph)	Conditions
WSW	8.9	Partly Cloudy

C) Actual user demand calculation:

In BSS, it often happens that a user returns the bike immediately after borrowing it at the same station, after which the user often borrows another bike. This phenomenon may be due to the user's dissatisfaction with the chair height or the current status of the bike. Therefore, if we directly count the number of records, the calculated user demand will be greater than the actual user demand.

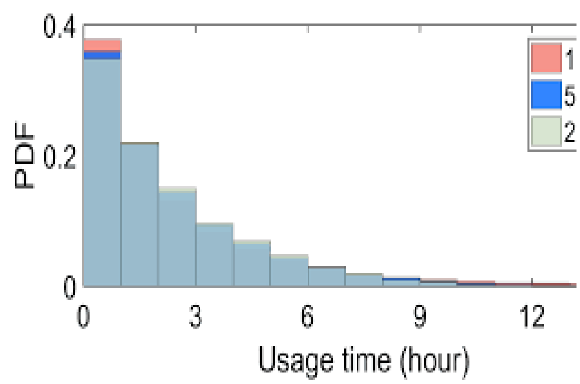

Fig 1: PDF of trip duration which begins and ends at the

same station.

As shown in Figure 1, the PDF curve of trip duration which begins and ends at the same station can be divided into a distinct spike and a long tail: for those real users, users at different stations could have different travel purpose, so the trip duration must be different. Due to the superimposed effects of records from all the stations, the travel time will be evenly distributed. Accordingly, the curve has a longer tail; and for the users who return the bike immediately, the trip duration is almost the same in each station, which leads to that very high spike. In the figure, the horizontal axis represents the riding time in seconds. The peak caused by the superimposed effects disappears at 120s. Therefore, records with trip duration less than 120s are treated as false records, and thus can be deleted from the original data. Finally, the actual demand can be calculated by simply accumulating the data in a half-hour unit.

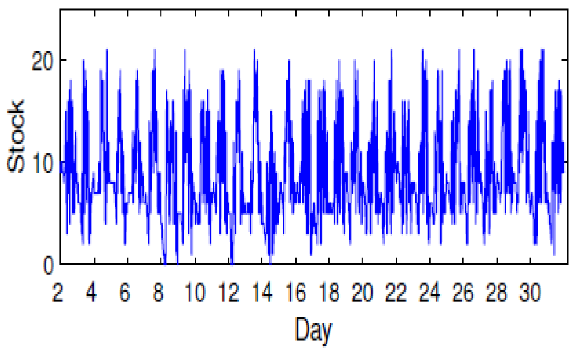

Fig. 3. Stock curve of station

presents CDF of empty and full hours in a month. Axis x represents full or empty hours. The calculation steps are as follows. Every half an hour, the inventory in each station is sampled once. There are 1440 samples in a month. When the inventory is less than 10% in these

TABLE 3 statistics on trip durations

< 15 min	15-30 min	30-45 min
53%	27%	11%
40-60 min	> 60 min	mean
5%	4%	23.31 min

samples, we consider it the empty time. Similarly, if the inventory is more than 90%, we consider it the full time. This picture can be used to measure the service level of the current bike-sharing system. It can be seen that about 19% of stations have an empty status for more than 200 hours in a month, 27% of the stations have been full for more than 200 hours in a month. The empty condition appears relatively less, and the full condition appears relatively more.

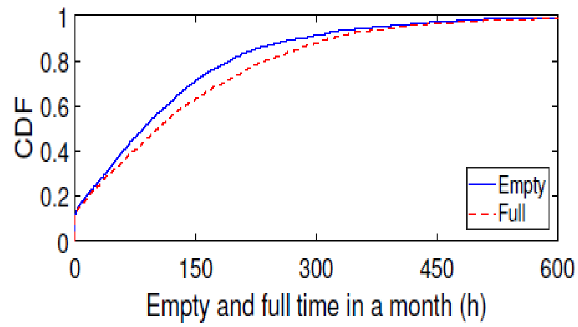

Fig. 7. CDF of empty and full hours

D) Other Objective Functions:

In practical applications, the advisor enables system operators to design other objective functions, thus achieving flexible resource scheduling. For example, we could advise users to rent bikes from active stations and still return them to active stations. Therefore, the aging process of a small part of bikes will be accelerated, allowing the regular upgrades of bikes in the system. Otherwise, it's unacceptable to the normal operation of the systems that many bikes need replacing at the same time.

Algorithm:

Input Start time t_0,user u, daily budget B, number of trips in each time interval $\{N_0{}', \cdots, N_k, \cdots, N_h \cdots\}$, number of total trips N, available prices offered $\{p_{0,\cdots,}p_{m\,,\cdots,}p_{q,\cdots}\}$;

Output: price r^n at iteration n;

1: Initialization:

- First time interval.$n = 0; h^o = h(t_o)$;
- Budgets. $B_k = \frac{N_{k.}\ B}{N}, \forall k \in (0, h); B = B_{k^n}, B^n = B$;
- Value estimates. $N_{u,m}^n, = 0, F_{u,m}^n = 0, \forall m \in [0, q]$;

2: for each request at time t do

3: if $k^n \neq k(t)$ then

4: $k^n = k(t)$;

5: $B^n = B^n + B_{k^n}, B = B^n$;

6: end if

7: $\widetilde{f_{u,m}^n} = F_{n,m}^n + \sqrt{\frac{21n(n)}{\dot{N}_{u,m}^n}}$;

8: $m\hat{}n = a\gamma g \max\{\min(\widetilde{F_{u,m}^n}, \frac{B}{N \cdot Pm})\}$s.t. $p_m \leq B^n$; $m \in [0, q]$

9: return $\gamma^n = Pm^n \cdot [dist(l_0, s_i) + dist(s_j, l_d) - min\{dist(l_o, s_i) + dist(s_j, l_d))] + f \cdot diSt(S_i S_j) - \min(dis\,t(s_i, s_j))]$;

10: end for

11: Feedback: Observe acceptance decision y^n;

12: Update Variables:

- $B^{n+1} = B^n - \gamma^n \cdot y^n;\ F^{n+1}_{u,m^n} = F^n_{u,m^n} + \frac{y^n - F^n_{u,m^n}}{N^n_{u,m^n}+1}$;
- $N^{n+1}_{u,m^n} = N^n_{u,m}n + 1;\ k^{n+1} = k^n\ ; n = n+1$

III. EXISTING SYSTEM

At present, operators perform bike redistribution based on monitor video and user complaints. However, this method has exposed the serious lag. It is usually when service unavailable events occur that operators start to give some scheduling instructions. When the vehicle arrives, service unavailable events may have passed for some time, which makes it difficult to meet the needs of users at rush hour. To increase service availability and enhance user experience, studies have been inducted to improve these bike redistribution strategies based on bicycle mobility models and predictions.

Disadvantages:

Most of the previous work focuses on bike usage patterns and rental volume forecasts for each station without considering online information Less attention has been devoted to demand prediction of each cluster from the view of bike flow mobility patterns which may not fit for recommending stations for users. In conclusion, developing a fine-grained prediction model involving multiple factors has proven to be elusive, and has remained a largely unstudied problem.

IV. PROPOSED SYSTEM

In this paper we propose a novel utilization aware trip advisor to lead users to help balancing bike usage without compromising the quality of service. We explore the overall characteristics of bike-sharing systems, analyze the spatial temporal patterns of user behavior and study the bike usage frequency, thus laying the foundation for trip advisor design. We introduce the concept of activeness to link bike usage frequency to station property which utilizes the topological characteristics of bike sharing network and the relative check out amount of each station. Meanwhile, we dynamically update the activeness to take the effect of the advisor on the system into account. We present a novel framework to balance bike usage with the help of users and validate our proposed method with real-world human mobility datasets.

A) *Advantages:*

To alleviate the unbalanced demand problem, we establish a fine-grained demand forecasting model and predict check in and check out demand on a per-station basis with sub-hour granularity by using random forest algorithm.

Starting from ensuring users' success rate of rental and return, the advisor is designed to dynamically recommend the optimal stations based on their current activeness of bike usage.

V. ARCHITECTURE & SYSTEM COMPONENTS

The below given framework depicts that users can send a query including their origin, destination and leaving time to the trip advisor and then get the recommended stations for rental and return. The key problem is how to guide the users to balance bike usage through station recommendation while not affecting the user experience. In this section, we will introduce the framework of our method, as shown in Figure. The framework is comprised of two major components: probabilistic forecasts and activeness calculation.

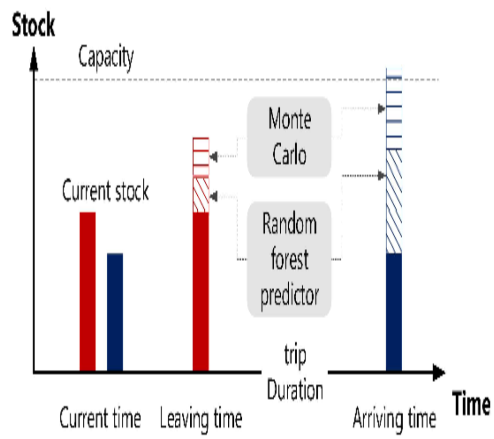

Fig 3: The idea of probabilistic forecasts.

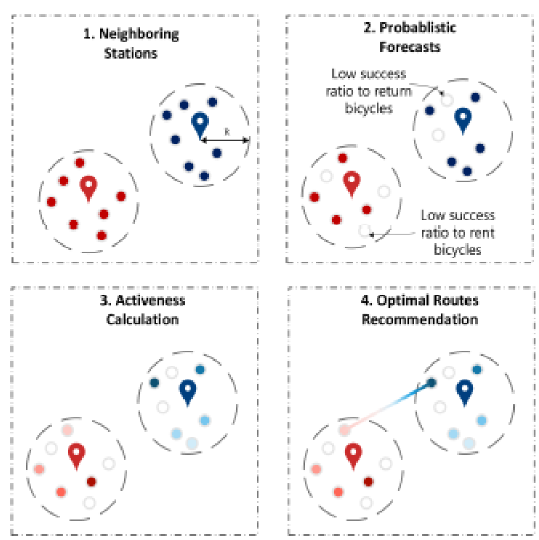

Fig: System Architecture-Framework of the trip advisor

A. *PROBABILISTIC FORECASTS:*

In order to encourage users to use the advisor and continue to help balancing bike usage, we first need to make sure that users can rent or return bikes successfully. Therefore, the first component, probabilistic forecasts, is designed to solve the no-service problem and guarantee the higher success rate for rental and return when users arrive at the stations. No-service means the situations in which a user can't find available bikes to rent, and those in which he/she finds there's no parking spot to return. This problem is mainly caused by the asymmetric and fluctuating user demand among the stations. For users, they may know where the nearest station is, but what they really want to know is the probability of successfully renting or returning bikes when he/she arrives there. To obtain the success rate at a precise moment, simply predicting the forthcoming user demand on half-hour granularity is not enough to meet the above requirement. The component of probabilistic forecasts is needed to predict the stock level on a minute timescale and further derive success rate through the Monte Carlo method. In Fig:3 the process is given by, at the beginning, the stock levels of candidate stations near the origin/destination are known. The forecasts consist of two parts. The first part is coarse-grained prediction using ran-doom forest model; the second part is fine-grained prediction based on Monte Carlo method. Here, we take predicting return success rate at arriving time as an example to elaborate on the details. Let [t] represent the rounded time of t to the nearest 30 minutes before. At the rounded current time [now], we already know the stock status rid of station i within R meters of the destination. Firstly, we predict the base check in and check out demand at each station with sub-hour granularity by using random forest model.

B. *Activeness Calculation:*

According to the previous analysis, active stations are characterized by the following properties:

1. Bikes returned to this station are easily checked out and flow to many other stations;
2. The stations that those bikes flowed to are also very active.

These properties remind us of the way to measure a web page's importance. PageRank is an algorithm used by Google Search to rank websites in their search engine results [16]. According to Google: PageRank works by evaluating the quality and quantity of links to a web page to determine a relative score of that page's importance. The idea that PageRank brought up is that more important websites are likely to receive more links from other websites. In bike-sharing systems, activeness can be defined to measure the active level of bike usage for each station based on the idea of PageRank. We begin by picturing the station network as a directed graph, with nodes represented by stations and edges represented by the bike flow (rent to return) between them. The underlying assumption is that more active stations in the network are likely to send more links to other stations. This makes sense because according to the analysis in Section 3, bikes do tend to be checked out extensively to many other stations at active stations and the bike usage in stations with more links out are usually more frequent. But this is only a start: the bikes must continue to flow to active stations so they can enter a high-speed circulation and be repeatedly used. This leads to the next assumption that stations which are themselves active weigh more heavily and help to make the stations that link to them active. If bikes rent from one station to stations with lower activeness, the bikes are likely to stay there, and it will take a long time for them to be checked out again. Therefore, this station may have low activeness as well.

B. *Admin:*

Admin module will give permission for user's registration and add vehicles and assign vehicles at each station based on peak timings suggest station services on neighboring stations requirement. Analyze user and station data at different times and suggest trip advisor. Admin will calculate activeness of using bikes at different time periods. Probabilistic Forecasts In order to encourage users to use the advisor and continue to help balancing bike usage

C. *User:*

There are two users in this system normal user and registered user. Normal user is the customer who want to search for bikes while traveling and find list of bikes available at wanted stations and at what time there is chance of getting data (probalistic forcasting and optimizing routes). Register user is the person who rents bikes from a station by giving his details like Aadhar card, driving license, start time, location, end point. This user will get information about where to give bike based on user requirement.

D.Station Services:

Station service module will handle each user details by registering each user Register user is the person who rents bikes from a station by giving his details like Aadhar card, driving license, start time, location, end point. This user will get information about where to give bike based on user requirement. Station services will get updated form admin about what are peak times and at which station at what time bikes should be move.

VI. CONCLUSION

In this paper, considering the investigation of general attributes, spatial transient examples and bike use in bike sharing, we propose a novel design of a use mindful use ware trip counsel which connects with clients to adjust bicycle use and delay the support interims of bicycles. Beginning from guaranteeing clients' prosperity rate of rental and return, the counselor is intended to powerfully prescribe the ideal stations dependent on their present liveliness of bicycle usage. We assessed the proposed framework through broad reenactments utilizing verifiable records from the world's biggest bicycle sharing framework, affirming the adequacy of our system.

VII. REFERENCES

[1] P. DeMaio, "Bike-sharing: History, Impacts, Models of Provision, and Future," Journal of Public Transportation, vol. 12, no. DeMaio 2004, pp. 41–56, 2009. [Online]. Available:http://www.transitinformatics.org/test/nctr/wpcontent/uploads/2010/03/JPT124DeMaio.pdf

[2] P. Midgley, "Bicycle-sharing schemes: enhancing sustainable mobility in urban areas," United Nations, Department of Economic and Social Affairs, pp. 1–12, 2011.

[3] L. Metro Bike, "2016 Year-end wrap-up will appear at the end of January," http://bike-sharing.blogspot.com/2017/01/2016-year-end-wrap-up-will-appear-at.html.

[4] J. Froehlich, J. Neumann, and N. Oliver, "Sensing and Predicting the Pulse of the City through Shared Bicycling," in IJCAI, 2009.

[5] A. Kaltenbrunner, R. Meza, J. Grivolla, J. Codina, and R. Banchs, "Urban Cycles and Mobility Patterns: Exploring and PredictingTrends in a Bicycle-based Public Transport System," Pervasive and Mobile Computing, vol. 6, no. 4, pp. 455–466, 2010.

[6] P. Vogel and D. C. Mattfeld, "Strategic and Operational Planning of Bike-Sharing Systems by Data Mining - A Case Study," inComputational Logistics, 2011, pp. 127–141.

[7] P. Borgnat, E. Fleury, C. Robardet, and A. Scherrer, "Spatial Analysis of Dynamic Movements of V´elo'v, Lyon's Shared Bicycle Program," in European Conference on Complex Systems (ECCS), 2009.

[8] Y. Li, Y. Zheng, H. Zhang, and L. Chen, "Traffic Prediction in a Bike Sharing System," in ACM SIGSPATIAL, 2015.

[9] Z. Online, "Hangzhou will add three public bicycle maintenance bases," http://zjnews.zjol.com.cn/system/2013/09/23/019608857.shtml.

[10] MSA, "BIKE SHARE PROGRAM," http://www.michigansuburbsalliance.org/wp-content/uploads/MSA-CAP-Strategy-Library Bike-Share-Program 20211017.pdf.

[11] Wikipedia, "Hangzhou Public Bicycle," https://en.wikipedia.org/wiki/Hangzhou Public Bicycle.

[12] S. Shaheen, H. Zhang, E. Martin, and S. Guzman, "China's hangzhou public bicycle: Understanding early adoption and behavioralresponse to bikesharing," Transportation Research Record: Journal of the Transportation Research Board, no. 2247, pp. 33–41,2011.

[13] dayoo, "Difficulty of renting public bicycles due to many damaged bikes," http://hainan.ifeng.com/news/detail 2014 06/18/2451605 0.shtml.

[14] J. L. Gastwirth, "A general definition of the lorenz curve," Econometrica: Journal of the Econometric Society, pp. 1037–1039, 1971.

[15] Z. Yang, J. Hu, Y. Shu, P. Cheng, J. Chen, and T. Moscibroda, "Mobility modeling and prediction in bike-sharing systems," in Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services, ser. MobiSys '16. New York, NY, USA: ACM, 2016, pp. 165–178.[Online].Available: http://doi.acm.org/10.1145/2906388.2906408

[16] S. Brin and L. Page, "Reprint of: The anatomy of a large-scale hypertextual web search engine," Computer networks, vol. 56, no. 18, pp. 3825–3833, 2012.

[17] N. Gast, G. Massonnet, D. Reijsbergen, and M. Tribastone, "Probabilistic forecasts of bike-sharing systems for journey planning," in ACM CIKM, 2015.

[18] J. Pfrommer, J. Warrington, G. Schildbach, and M. Morari, "Dynamicvehicle redistribution and online price incentives in shared mobility systems," IEEE Transactions on Intelligent TransportationSystems, vol. 15, no. 4, pp. 1567–1578, 2014.

[19] C. Fricker and N. Gast, "Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity," Euro journal on transportation and logistics, vol. 5, no. 3, pp. 261–291, 2016.

[20] A. Singla, M. Santoni, G. Bart´ok, P. Mukerji, M. Meenen, and A. Krause, "Incentivizing users for balancing bike sharing systems." in AAAI, 2015, pp. 723–729.

[21] Panshi, "Hangzhou Micro Bus," http://www.myhz.com/?p=97.

[22] S. a. Shaheen, S. Guzman, and H. Zhang, "Bikesharing in Europe, the Americas, and Asia," Transportation Research Record: Journal of the Transportation Research Board, vol. 2143, pp. 159–167, 2010.

[23] S. a. Shaheen, A. P. Cohen, and E. W. Martin, "Public Bikesharing in North America: Early Operator Understandingand Emerging Trends," Transportation Research Record: Journal of the Transportation Research Board, vol. 2387, pp. 83–92, 2013. [Online]. Available: http://trb.metapress.com/openurl.asp?genre=articlef&gid=doi:10.3141/2387-10

[24] E. W. Martin and S. A. Shaheen, "Evaluating Public Transit Modal Shift Dynamics in Response to Bikesharing: A Tale of Two U.S. Cities," Journal of Transport Geography, vol. 41, pp. 315–324, 2014.[Online].Available: http://linkinghub.elsevier.com/retrieve/pii/S0966692314001409

[25] S. D. Parkes, G. Marsden, S. A. Shaheen, and A. P. Cohen, "Understanding the Diffusion of Public Bikesharing

Systems:Evidence from Europe and North America," Journal of Transport Geography, vol. 31, pp. 94–103, 2013. [Online]. Available:http://linkinghub.elsevier.com/retrieve/pii/S0966692313001130

[26] J. Liu, Q. Li, M. Qu, W. Chen, J. Yang, H. Xiong, H. Zhong, and Y. Fu, "Station site optimization in bike sharing systems," in Data Mining (ICDM), 2015 IEEE International Conference on. IEEE, 2015,pp. 883–888.

[27] X. Wang, C. Yuen, N. U. Hassan, N. An, and W. Wu, "Electric vehicle charging station placement for urban public bus systems," IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 1,pp. 128–139, 2017.

[28] E. O. Mahony and D. B. Shmoys, "Data Analysis and Optimization for (Citi) Bike Sharing," in AAAI, 2015.